Supplementary Appendix

Supplement to:

Dobbs K, Domínguez-Conde C, Zhang SY, Parolini S, Audry M, Chou J, Haapaniemi E, Keles S, Bilic I, Okada S, Massaad MJ, Rounioja S, Alwahadneh AM, Serwas NK, Capuder K, Ciftci E, Felgentreff K, Ohsumi T, Pedergnana V, Boisson B, Haskoloğlu S, Ensari A, Schuster M, Moretta A, Itan Y, Patrizi O, Rozenberg F, Lebon P, Saarela J, Knip M, Petrovski S, Goldstein DB, Parrott RE, Savas B, Schambach A, Tabellini G, Bock C, Chatila T, Comeau AM, Geha RS, Abel L, Buckley RH, Ikincioğullari A, Al-Herz W, Helminen M, Doğu F, Casanova JL, Boztuğ K, Notarangelo LD.

Human *DOCK2* mutations underlie an autosomal recessive pleiotropic immunodeficiency with early-onset invasive infections

Table of Contents

Supplementary Methods

Study subjects

The five index patients were from Lebanon (P1), Finland (P2), Turkey (P3 and P4) and Honduras/Nicaragua (P5). They were included with informed written consent by the parents and approval from the Institutional Review Boards of Kuwait Ministry of Health, the INSERM Institute, the Rockefeller University, the Medical University of Vienna, Ankara University Medical School, Boston Children's Hospital, and the Duke University Medical School.

Genetic studies

The underlying genetic defects in the *DOCK2* gene were identified by whole exome sequencing (WES), with or without additional positional information generated by genome-wide linkage analysis (P3) or homozygosity mapping (P4). Genomic DNA was extracted from whole blood using the Gentra Puregene kit (Qiagen) for P1 and P5, the FlexiGene DNA kit (Qiagen) for P2, the Qiamp midi DNA kit (Qiagen) for P3, the Wizard Genomic DNA Purification kit (Promega) for P4.

For P3, four members of family 3 were genotyped with the Affymetrix genome-wide single nucleotide polymorphism (SNP) 6.0 array. Genotype calling was achieved with Affymetrix Power Tools (http://www.affymetrix.com/partners_programs/programs/developer/tools/powertools.affx) for the four family members. We discarded monomorphic SNPs, SNPs with a call rate lower than 100% and SNPs presenting more than one Mendelian inconsistency in the family. SNPs were further filtered with population-based filters. We then used about 111,600 high-quality SNP markers to carry out linkage analysis, assuming autosomal recessive inheritance with complete penetrance. Parametric multipoint linkage analysis was carried out with Merlin[1]. The Turkish family founders and HapMap CEU trios were used to estimate allele frequencies and to define linkage clusters, with an r^2 threshold of 0.4. For P4, Affymetrix 6.0 SNP-based homozygosity mapping was performed to detect potential homozygous intervals. The protocol was carried out according to the Affymetrix® Genome-Wide Human SNP Nsp/Sty 6.0 protocol. Genotype calling was performed using the Affymetrix® Genotyping Console™ software. Detection and annotation of the homozygous

intervals was performed using HomozygosityMapper[2] and PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/)[3].

For WES, Illumina paired-end libraries were created using the SureSelect Human All Exon V5 kit (Agilent), the Nextera exome enrichment kit (Illumina) and the Nimblegen SeqCap EZ V4.0.

For patient P1, libraries were sequenced using the Illumina HiSeq1500 sequencing platform with 101 bp read length. Average coverage of the exome for P1 was 211x with 98% of the exome covered at a depth of ≥10x. The read mapping, variant calling and filtering steps for germ-line variants were performed as described previously[4].

For patients P2 and P3, captured libraries were sequenced with the Illumina HiSeq2500 and Illumina HiSeq2000 sequencing platforms, respectively, and reads were mapped to the human reference genome assembly GRCh37 (UCSC hg19) using the Burrows-Wheeler Aligner (BWA)[5]. Variants were identified with the Genome Analysis Toolkit (GATK)[6], SAM tools[7] and Picard Tools (http://broadinstitute.github.io/picard/). Variant calls with a read coverage ≤2x and a Phred-scaled SNP quality ≤20 were eliminated.

For patient P4, a multiplexed pool of twelve samples was sequenced on 4 lanes of the Illumina HiSeq2000 sequencing platform in 100 bp paired-end mode. Reads were aligned to GRCh37 using BWA, as indicated above. Duplicate reads were marked with Picard Tools (http://picard.sourceforge.net), while insertion/deletion realignment and base quality score recalibration was performed utilising the GATK. Variant calling was performed in the cohort of 12 samples using the GATK Unified Genotyper, before applying Variant Quality Score Recalibration (VQSR) (1000Genomes, HapMap, dbSNP131). All thresholds for GATK tools were based on the GATK Best Practice Variant Detection v3 recommendations[8].

For patient P5, sequencing was completed on the Illumina HiSeq2500 sequencing platform in 125bp paired-end mode using KAPA Biosystem's library preparation kit followed by whole exome capture using Nimblegen SeqCap EZ V4.0. After processing the raw reads using a pipeline based on the GATK best practice protocol, the resulting alignments had ≥10-fold coverage in over 97.4% of the 33,266,994 bases of consensus coding sequence CCDS (release 14). The sequence data were analyzed using established protocols that identify qualifying variants forming genotypes not observed in a large control database of up to 61,486 samples provided by the Exome Aggregation Consortium (ExAC).

For all patients, candidate mutations were verified by capillary sequencing using BigDye chemistry (Life Technologies).

Sequence analysis

We used the SIFT[9], PolyPhen-2[10] and CADD[11] algorithms to assess the pathogenicity of the identified DOCK2 missense mutations.

T Cell Receptor Excision Circle (TREC) Assay

Dried blood spots collected at birth from P2 and P4 were retrieved upon informed consent. Real-time quantitative multiplex PCR was used to measure levels of TREC and RNaseP (as internal control) in template genomic DNA, as previously described[12].

Generation of T and NK cell lines

Peripheral blood mononuclear cells (PBMCs) were separated by Ficoll density gradient. For the generation of T cell lines, cells were activated and expanded *in vitro* using the human T cell activation/expansion kit according to manufacturer's instructions (Miltenyi). To obtain polyclonal NK cell lines, NK lymphocytes were isolated from PBMCs using negative selection (NK cell isolation Kit, Miltenyi), then cultured on irradiated feeder cells in the presence of 100 U/mL recombinant human IL-2 (Proleukin; Chiron) and 1.5 ng/mL phytohemagglutinin (GIBCO Ltd).

Western Blot

Whole cell lysates were prepared from 10^6 cells from control- and patient-derived T cell lines, loaded on 8-16% polyacrylamide TRIS-HEPES-SDS gel (Pierce Laboratories) and separated by SDS-PAGE. Proteins were transferred to polyvinylidene difluoride (PVDF) membrane using iBlot system (Invitrogen). Blots were probed overnight with 1:400 dilution of rabbit polyclonal anti-human DOCK2 (N-terminus) antiserum (AbCam), and 1:10,000 dilution of anti-β-Actin monoclonal antibody (Cell Signaling) as a loading control. Bands were revealed using SuperSignal West Femto kit (Fisher).

In a separate set of experiments, whole cell lysates were prepared from 10^6 cells from control- and patient-derived EBV-immortalized B cell lines, or SV40-immoralized

fibroblasts with or without stable expression of exogenous wild type DOCK2 expression through lentiviral transduction. Equal amounts of protein from each sample were separated by SDS-PAGE and blotted onto PVDF membrane (Bio-Rad, Hercules, CA), which was then probed with a goat anti-human DOCK2 antibody (R&D Systems, Minneapolis, MN), and with anti-GAPDH (Santa Cruz, California, CA) or anti-β-tubulin **(Sigma-Aldrich, St.Louis,** MO) antibodies as loading controls. Antibody binding was detected by enhanced chemiluminescence (ECL; Amersham-Pharmacia-Biotech).

GTP-bound RAC1 pull-down assay

To evaluate TCR-induced RAC1 activation, 10^7 cells of patients and control-derived T cell lines were rested in culture medium overnight, then activated with or without 1 mg/ml anti-CD3 mAb (Miltenyi) for 30 min at 37°C. The cell lysates were then incubated with 20μg of Glutathione S transferase and GTPase binding domain of p21-activated kinase (GST-Pak1-PBD), immunoprecipitated with glutathione resin according to manufacturer's instructions (Pierce laboratories), and immunoblotted using anti-RAC1 monoclonal antibody (Pierce laboratories). Aliquots of the cell extracts were immunoblotted in parallel as loading controls.

Flow cytometry

For analysis of cell surface markers, $2x10^6$ PBMCs were used as starting material or 100 ml of whole blood in EDTA was lyzed using RBC Lysis Buffer (eBioscience). Cells were washed twice in FACS buffer (PBS with 2% FBS), and resuspended in 100 μl of FACS buffer with antibodies for 30 min on ice. Cells were washed twice in FACS buffer, resuspended in 300 μl of FACS buffer. The following monoclonal antibodies (mAbs) were used in various combinations for phenotypic analysis of T and B lymphocytes: FITC-anti-CD3, PE-anti-CD4, PE-anti-CD16, APC-anti-CD8, APC-anti-CD56, APC-anti-CD31, APC-anti-CD24, PerCP Cy5.5-anti-CD19, PerCP Cy5.5-anti-CD3, PerCP Cy5.5-anti-CD45RA, PerCP Cy5.5-anti-CD45R0, PE Cy7-anti-CD3, Pacific Blue-anti-CCR7 (all from Biolegend), FITC-anti-CD45RA, FITC-anti-IgD, PE-anti-CD21, V450-CD8, V450-CD56, AF700-CD45RA, PE-Cy7-CD19 and APC-H7-CD3 (all from BD Biosciences), FITC-anti-CD57, FITC-anti-CD38, APC-anti-CD27, PerCP Cy5.5-anti-CD45, APC-anti-CD3, APC-anti-CD3 (all from

eBioscience), PE-anti-IgM (from Southern Biotech), and PE-anti-CD4 and FITC-anti-CD16 (both from Beckman Coulter). Cells were analyzed on a LSR Fortessa (BD Biosciences). Data were analyzed with FlowJo software version 8.8.7 or version 10 (TreeStar).

For phenotypic analysis of NK lymphocytes, cells were first stained with appropriate primary mAbs, which included: anti-NKp46 (BAB281), anti-NKp30 (AZ20), anti-NKp44 (Z321 and AZ140), anti-NKG2D (ON72, BAT221, and ECM217), anti-CD16 (SUS142), anti-KIR2DL1/S1 (11PB6), anti-KIR2DL2/L3/S2 (GL183), anti-KIR3DL1/S1/L2 (AZ158), anti-NKG2A (Z199), anti-CD57 (XA147). Cells were then stained with PE-conjugated isotype-specific goat-anti-mouse antiserum (Southern Biotechnology, Birmingham, AL), followed by staining with a mixture of FITC-labeled anti-CD3, FITC-anti-CD14, FITC-anti-CD20 and PC5-anti-CD56. Analysis of NK cell markers was performed by gating on $CD56^{+}CD3^{-}CD14^{-}CD20^{-}$ cells. Cells were analyzed on a FACSCanto (BD Biosciences). Data were analyzed with Diva v6.1.3 software (BD Biosciences, Mountain View, CA). All primary mAbs used for NK cell phenotyping were generated at the Department of Molecular and Translational Medicine, University of Brescia, Italy and in the Laboratory of Molecular Immunology, DIMES, University of Genoa, Italy, and have been previously described[13].

For the analysis of NKT lymphocytes, PBMCs were stained with PE-Texas red or FITC- conjugated CD45, APC/Cy7 or PerCP5.5-conjugated CD3, FITC or PECy7-conjugated CD4, APC or APC/Cy7-conjugated CD8, and PE or APC-conjugated, PBS57-loaded, CD1d-tetramers (with unloaded CD1d tetramers used as control). Upon gating on $CD45^{+}$ cells, NKT cells were identified as $tetramer^{+}$ $CD3^{+}$. Events were collected on a BD Canto flow cytometer (BD Biosciences) and analyzed with FlowJo software, version 8.8.7. CD1d-tetramers were obtained from National Institute of Allergy and Infectious Disease Major Histocompatibility Complex Tetramer Core Facility, Atlanta, GA.

Analysis of NK cell degranulation and IFN-γ production.

For degranulation assay against erythroleukemia human cell line K562 and mastocytoma murine cell line P815, PBMCs derived from patients P1 and P2 and from healthy donors were incubated with or without 100 U/mL recombinant human

IL-2 (Proleukin; Chiron) at 37°C overnight (in the case of K562 target cell assay) or for three days (for P815 target cell assay). Cells were then incubated with target cells at an effector:target ratio of 1:2 (for K562 cells) and 1:7 (for P815 cells), in a final volume of 200 µl in round-bottomed 96-well plates at 37°C and 5% CO2 for 4 hours in culture medium supplemented with anti-CD107a-PE monoclonal antibody. Degranulation assay was performed in the absence or presence of various mAbs including anti-NKp30, anti-NKp46, anti-Nkp44, anti-CD16, anti-NKG2C and anti-NKG2D, respectively. After 1 h of co-incubation, GolgiStop (BD Biosciences Pharmingen, San Diego, CA, USA) was added at a 1:100 dilution. Surface staining was done by incubating the cells with anti-CD3, anti-CD14, anti-CD20 and anti-CD56 mAbs for 30 min at 4°C. The cells were washed and analyzed by flow cytometry (FACSCanto, BD Biosciences). Analysis of NK cell degranulation was performed by gating on $CD56^{+}CD3^{-}CD14^{-}CD20^{-}$ cells and analyzing the proportion of cells expressing CD107a on the cell surface. The re-directed killing assay against P815 target cells was also performed using polyclonal NK cells that had been grown *in vitro* as described above.

To detect intracellular production of IFN-γ, PBMCs derived from patient P2 and from healthy donors were incubated overnight at 37°C with IL-12 (0.5 ng/ml), or IL-12 (0.5 ng/ml) and IL-18 (0.1 ng/ml) combined. Cells were then washed, fixed and permeabilized with BD Cytofix/Cytoperm kit (BD Biosciences Pharmingen). IFN-γ production was detected by subsequent intracellular staining with anti-IFN-γ-PE (BD Biosciences Pharmingen) upon gating on $CD56^{+}CD3^{-}CD14^{-}CD20^{-}$ cells.

For both degranulation and IFN-g expression experiments, the percentage of positive cells was calculated subtracting the baseline CD107a or IFN-γ expression in control cultures in the absence of stimuli (target cells or cytokines).

Chemotaxis assay

$5x10^{5}$ PBMCs from patients and controls were suspended in 100 µL RPMI media with 10% FBS. The cells were loaded into 5.0 µm Transwell inserts (Corning) which were placed onto 24-well plates containing 450 mL of RPMI medium alone or supplemented with 1.5µg/mL CCL21 or with 800ng/mL CXCL12 (Peprotech). After 3 hours of incubation at 37°C, cells that had migrated to the lower chamber were collected, stained with PerCP-Cy5.5 labeled anti-CD19 and with APC-conjugated

anti-CD3 (both from Biolegend) and counted by flow cytometry (BD Fortessa). Results were calculated as percentage of cells that had migrated out of total input cells.

F-Actin Polymerization

For analysis of chemokine-induced actin polymerization in T and B lymphocytes, PBMCs ($2.5x10^6$ cells/mL in RPMI, 1% FCS) were allowed to rest at least 3 hours on ice, then stimulated with 800 ng/ml CXCL12. Aliquots were taken at the indicated time points (between 0 and 60 seconds), and immediately fixed in 4% paraformaldehyde for 10 min. After washing with PBS and FcgR blocking with human serum (1:10 dilution in PBS), cells stained with FITC-conjugated anti-CD3 and PerCP-Cy5.5 labeled anti-CD19 mAb (BioLegend), permeabilized using BD Cytofix/Cytoperm Kit, stained with AlexaFluor 647-conjugated phalloidin (Molecular Probes), and analyzed by flow cytometry (BD Fortessa) using FlowJo software version 8.8.7 (TreeStar).

For analysis of actin polymerization in NK lymphocytes, 10^5 polyclonal IL2-activated NK cells/tube were re-suspended in 100 µl RPMI media and rested overnight in RPMI without IL-2. After 30 min incubation on ice, cells were incubated with CD16- and NKp46-specific primary antibodies on ice for one hour, then warmed at 37°C for 2 min, incubated and cross-linked with isotype-specific secondary antibodies for 30s and fixed with 4% paraformaldehyde for 10 min. After washing with PBS, cells were re-suspended in 200 µl of Fix/Perm solution (BD Biosciences) at 4°C for 20 min, washed twice with 0.1% saponin solution, stained with CF^{TM}633 Dye-conjugated phalloidin (Biotium) (1:400 dilution in saponin) at 4°C for 30 minutes and then re-suspended in 500 µl of PBS solution for flow cytometry analysis. Events were acquired on BD Fortessa, and analyzed with FlowJo software version 8.8.7 (TreeStar). Results were expressed as phalloidin mean fluorescence intensity (MFI) at each time point.

Flow cytometry analysis of ERK1/2 and MEK1 phosphorylation

10^5 polyclonal IL2-activated NK cells/tube were re-suspended in 100 µl RPMI media, rested for 48 hours in RPMI without IL-2, and incubated at 4°C for 15 min. Cells were then incubated for 1 hour at 4°C with various primary antibodies specific for CD16, NKp30 and NKp46. After cross-linking stimulation with isotypic specific

secondary antibodies for 3-5 min, the cells were incubated with 200 µl of Fix/Perm solution (BD Biosciences) at 4°C for 20 min, washed twice with 0.1% saponin solution, then stained for 30 minutes at 4°C with Alexa Fluor 488 Mouse Anti-ERK 1/2 – pT202/pY204 or with Alexa Fluor 488 Mouse Anti-MEK 1 – pS298 (both from BD Biosciences), and re-suspended in 500 µl of PBS. Events were acquired on BD Fortessa, and analyzed with FlowJo software version 8.8.7 (TreeStar).

Cytokine production

To measure cytokine production, PBMCs isolated by Ficoll-Paque density gradient centrifugation were stimulated with intact viruses for 24 hours, at a concentration of 2×10^6 cells/ml in RPMI 1640 supplemented with 10% FCS. We used herpes simplex virus-1 (HSV-1, strain KOS-1, multiplicity of infection (MOI)=1), and vesicular stomatitis virus (VSV, strain Indiana, MOI=1). The production of IFN-α, -β, -λ and IL-6 was assessed by ELISA. ELISA for IFN-α (AbCys SA, Paris, France), IFN-β (TFB, Fujirebio, Inc., Tokyo, Japan), and IL-6 (Sanquin, Amsterdam), respectively, was carried out according to the kit manufacturer's instructions. The IFN-λ ELISA was developed in the laboratory. Briefly, plates were coated by incubation overnight at 4°C with 1 µg/ml of anti-human IFN-λ1 monoclonal antibody (R&D Systems, Minneapolis, MN), and the IFN-λ1 concentration in the supernatant was measured using a secondary biotinylated anti-human IFN-λ1 monoclonal antibody (R&D Systems, Minneapolis, MN) at a concentration of 400 µg/ml.

Determination of mRNA levels by RT-qPCR

Total RNA was extracted from PBMCs. RNA was reverse-transcribed directly, with random hexamers, to determine mRNA levels for different subtypes of IFNA, IFNB, IFNL1, IFIT1, MX1 and IL6. Reverse transcription-quantitative PCR (RT-qPCR) was performed with Applied Biosystems Assays-on-Demand™ probe/primer combinations and 2 x universal reaction mixture, in an ABI PRISM® 7700 Sequence Detection System. We used β-glucuronidase (GUS) for normalization. Results are expressed according to the ΔΔCt method, as described by the manufacturer.

Viral replication in fibroblasts and effects on cell viability

10^5 SV40-fibroblasts were plated in individual wells of 24-well plates and infected with vesicular stomatitis virus (VSV), at a MOI of 0.1, in DMEM medium supplemented with 2% FCS. After 30 minutes, cells were washed and incubated in 500 µl of medium. Supernatants were obtained at the 2h, 8 h, 16h, 24h and 48 h time points and frozen. VSV titers were determined by calculating the 50% end point ($TCID_{50}$) after the inoculation of Vero cell cultures in 96-well plates.

The viability of SV40-fibroblasts upon VSV or EMCV infection was assessed by resazurin oxidoreduction (TOX-8) (Sigma-Aldrich, St. Louis, Missouri). Cells were plated in triplicates in 96-well flat-bottomed plates (2 x 10^4 cells/well) in DMEM supplemented with 10% FCS; 18 hours later, cells were infected by incubation for 24 hours with VSV at the indicated MOI. Resazurin dye solution was then added, at a volume one tenth that of the culture medium, and the samples were incubated for an additional two hours at 37°C. Fluorescence was then measured at a wavelength of 590 nm, using an excitation wavelength of 560 nm. Background fluorescence, calculated for dye and complete medium alone (in the absence of cells) was then subtracted from the values for all the other samples; 100% viability corresponds to the fluorescence of uninfected cells. For assays of cell protection upon viral stimulation, cells were treated with IFN-α2b (Schering-Plough, Brussels, Belgium, 1 x 10^3 IU/ml), starting the same time as infection

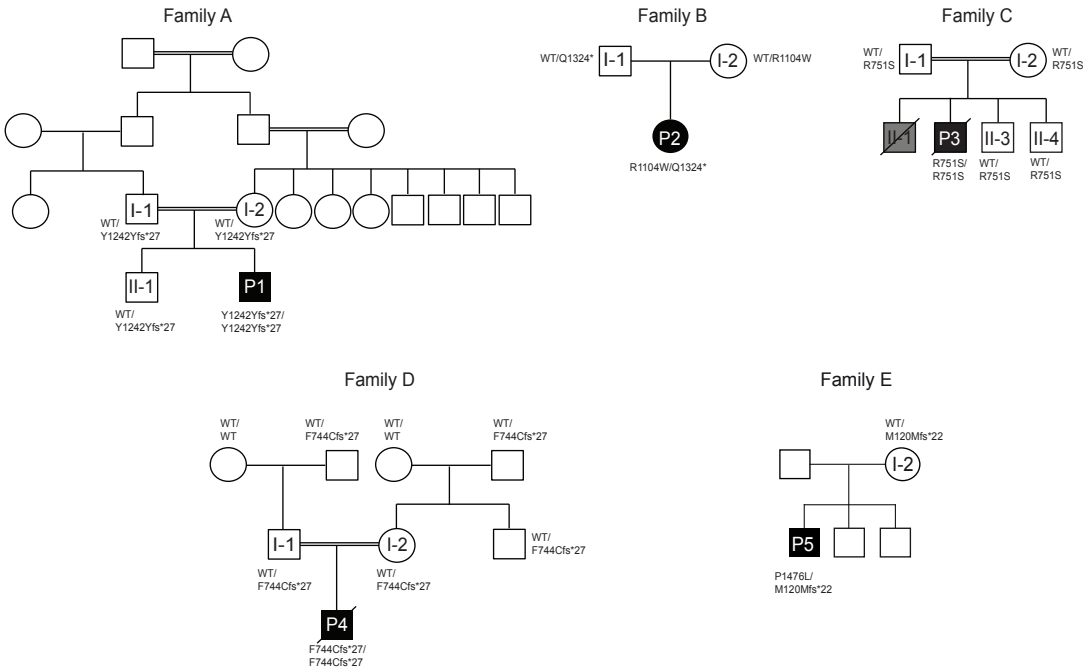


Figure S1. Segregation of *DOCK2* mutations. Extended patient pedigrees showing the segregation of the identified *DOCK2* mutations. In all pedigrees, the identified *DOCK2* mutations fully segregate with the disease phenotype under the assumption of autosomal recessive inheritance with full penetrance. Since no DNA was available from P5's father, the possibility of the p.P1476L mutation being a *de novo* mutation cannot be ruled out.

Figure S2

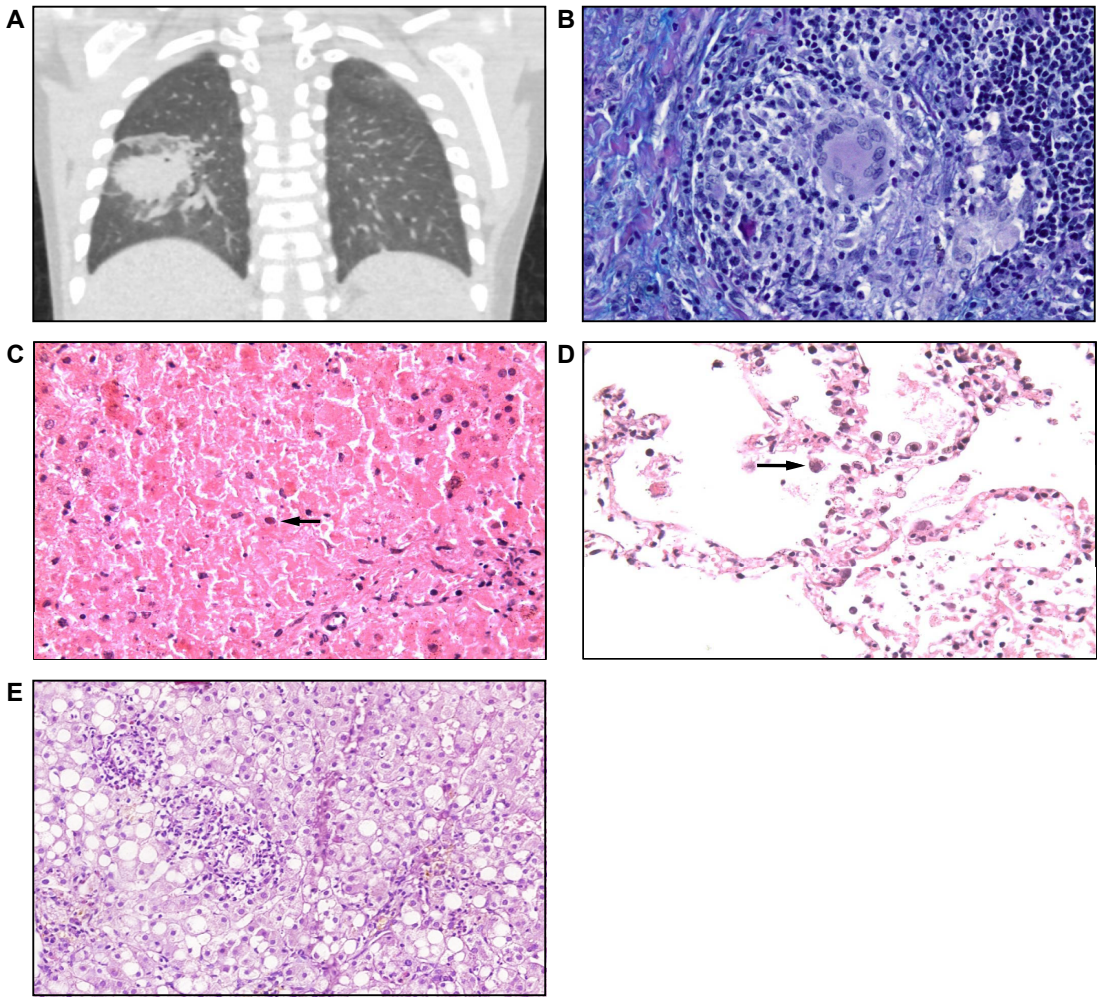


Figure S2. Clinical features identified in DOCK2-deficient patients. (A) CT scan of the lung showing granuloma in P2. (B) Lung immunopathology showing granuloma in P2. (C) Postmortem histology of a liver specimen from P3 showing a degraded hepatocyte with a nuclear inclusion (black arrow). The background liver parenchyma is necrotic. (D) Postmortem histology of lung specimen from P3 showing polymorphonuclear leukocytes and necrotic fragments in the alveolar spaces. A cell with a nuclear inclusion (black arrow) is indicated. In C and D: Hematoxylin and eosin staining; original magnification x400. (E) Liver biopsy showing significant hepatocyte macroversicular fatty change, lobular and portal inflammation with eosinophils, yellow-brown pigment accumulation in P4 (hematoxylin/eosin staining, magnification 200x).

Figure S3

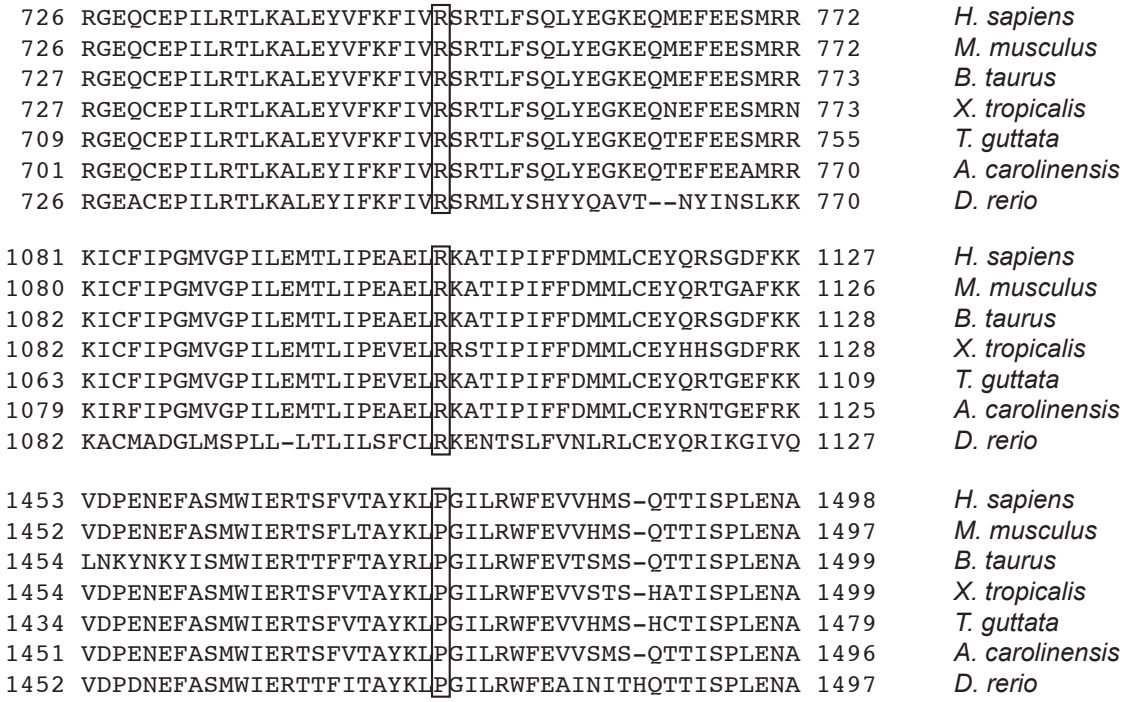


Figure S3. Evolutionary conservation of the identified *DOCK2* missense mutations. Amino-acid sequence of the region containing residues R751, R1104 and P1476 (all three indicated by a black frame) mutated in patients P3, P2 and P5 respectively, illustrating that the positions affected by the mutations are conserved throughout evolution.

Figure S4

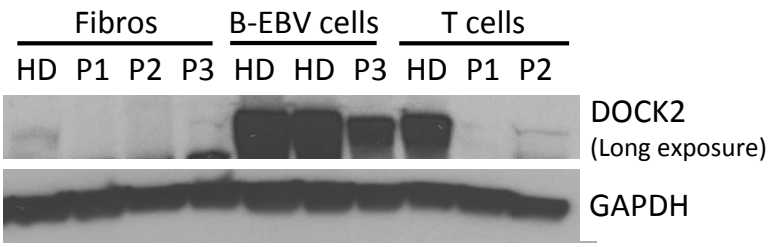


Figure S4. Western blot analysis of DOCK2 expression. Western-blot analysis of SV-40-immortalized primary fibroblasts derived from P1, P2 and P3, EBV-B cells derived from P3 and T cells derived from P1 and P2 showing absence or severely reduced expression of DOCK2 protein, in comparison with those from two healthy donors (HD). GAPDH was used as a loading control.

Figure S5

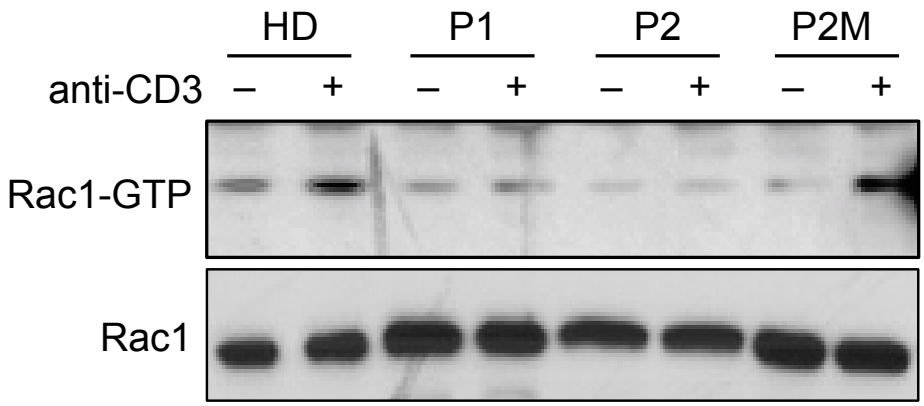


Figure S5. Rac1 activation pull-down. Defective Rac1 activation in T cell lines derived from P1 and P2 upon stimulation of the T cell receptor with anti-CD3 monoclonal antibodies. The figure displays the raw data used to generate the bar graph shown in Figure 2A.

Figure S6

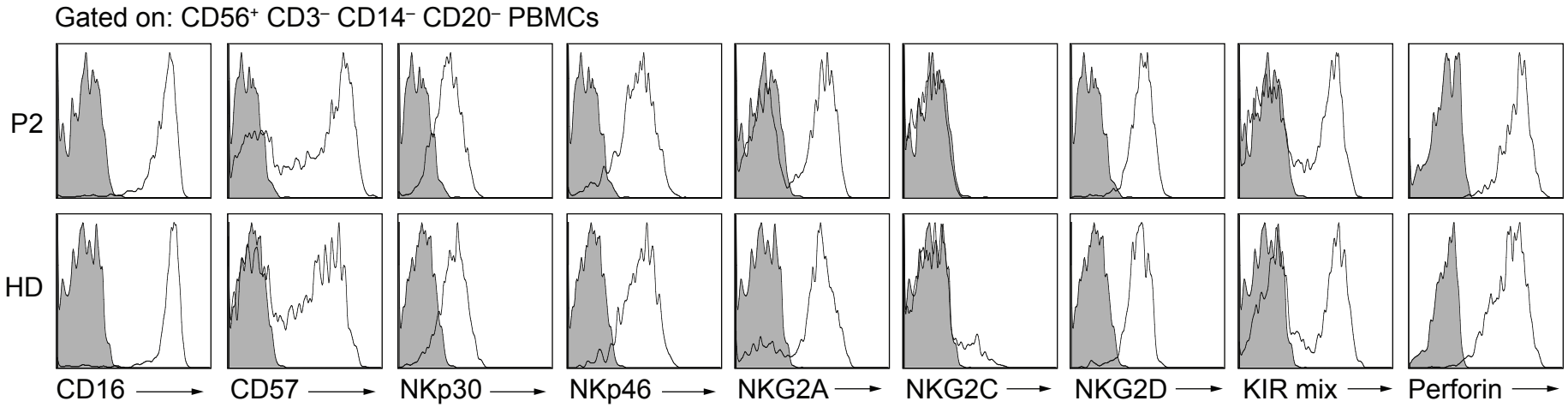


Figure S6. Phenotypic analysis of circulating NK-cells. Expression of cell surface molecules and intracellular perforin on $CD3^-$ $CD56^+$ NK cells analyzed by flow cytometry. KIR mix – a cocktail of monoclonal antibodies recognizing KIR2DL1, KIR2DL3 and KIR3DL2.

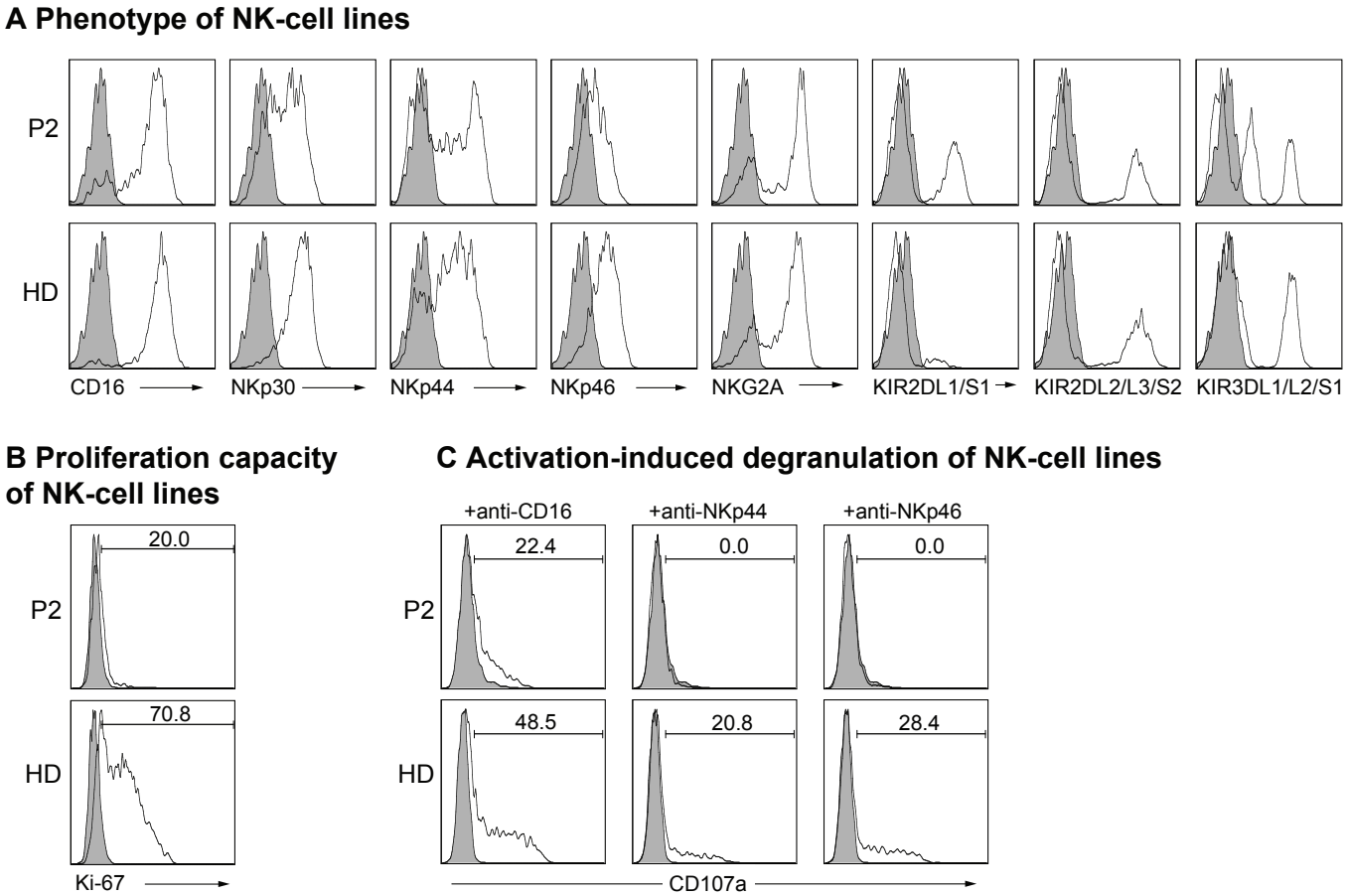


Figure S7. Phenotype, proliferation capacity and activation-induced degranulation of patient-derived IL-2-activated polyclonal NK-(100U/ml)cell lines. (A) Cell surface expression of NK-cell surface markers. (B) Representative flow cytometry histograms showing proliferation capacity of P2-derived NK-cell line cultured in the presence of IL-2. (C) Impaired CD107a expression on cell surface of P2-derived, IL-2-activated polyclonal NK-cell line upon engagement of CD16, NKp44 and NKp46, respectively. The numbers in the gates denote the increase in the percentage of $CD107a^+$ cells upon stimulation (solid line histograms) compared to unstimulated controls (grey histograms).

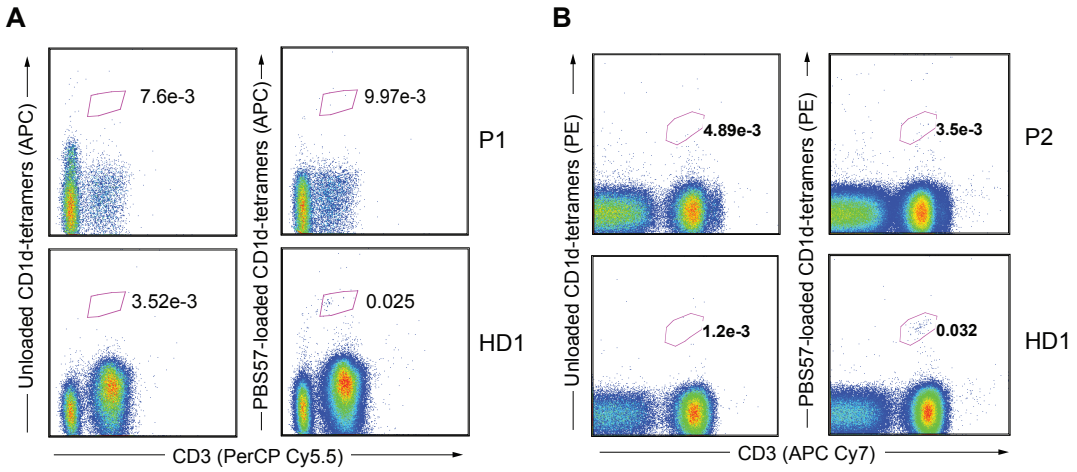


Figure S8. Identification of NKT cells in DOCK2-deficient patients. Flow cytometry analysis of CD1d-tetramer staining of $CD3^+$ PBMCs from patient P1 and P2, illustrating absence of NKT cells in the DOCK2-deficient patients.

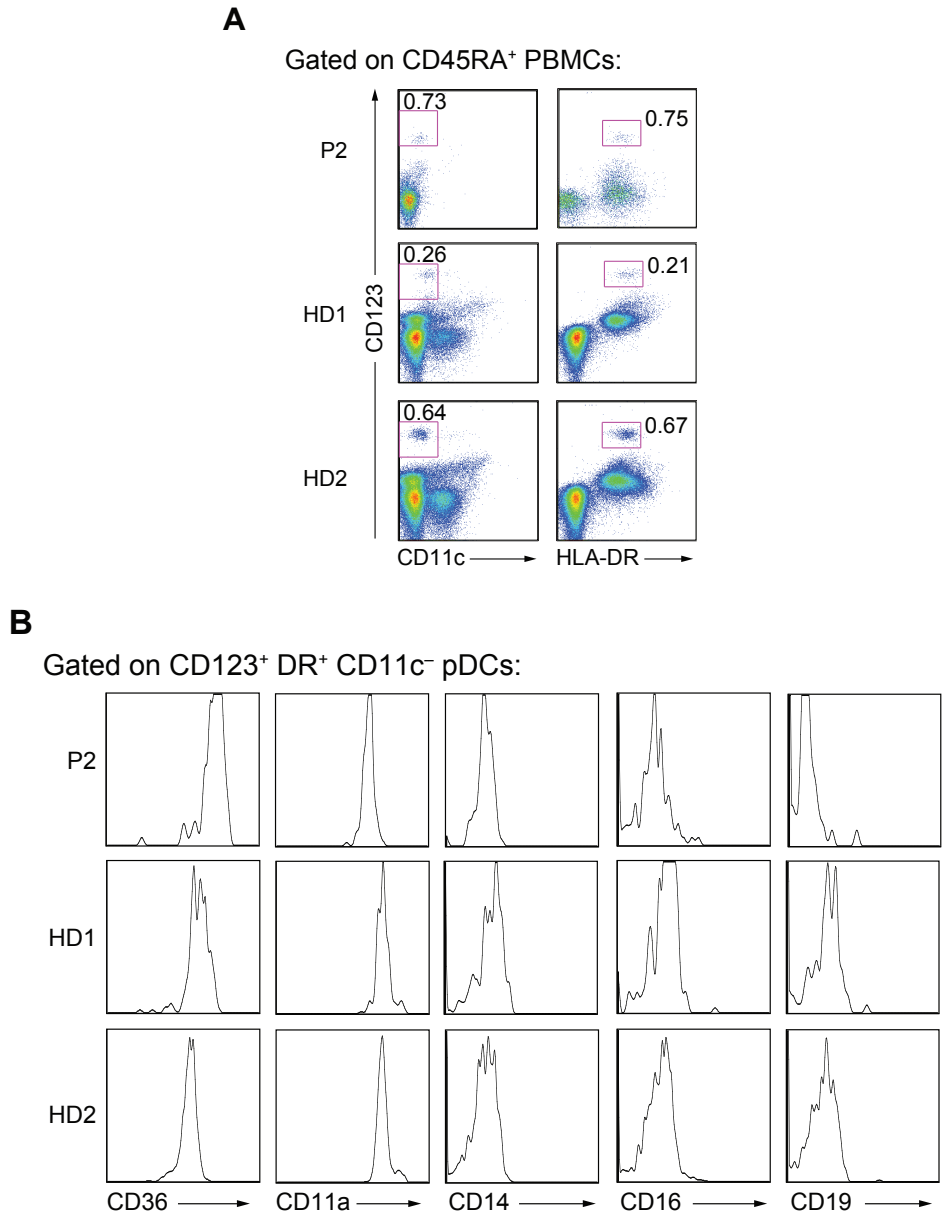


Figure S9. Identification and phenotypic characterization of pDCs in patient P2. (A) Flow cytometry dot-plots showing comparable relative numbers of $CD123^+$ $HLA\text{-}DR^+$ $CD11c^-$ pDCs (within the normal range) in PBMCs from patient P2 and healthy controls HD1 and HD2. (B) Unaltered cell-surface marker expression on pDCs from P2 and a healthy donor (HD), shown by the representative flow cytometry histograms.

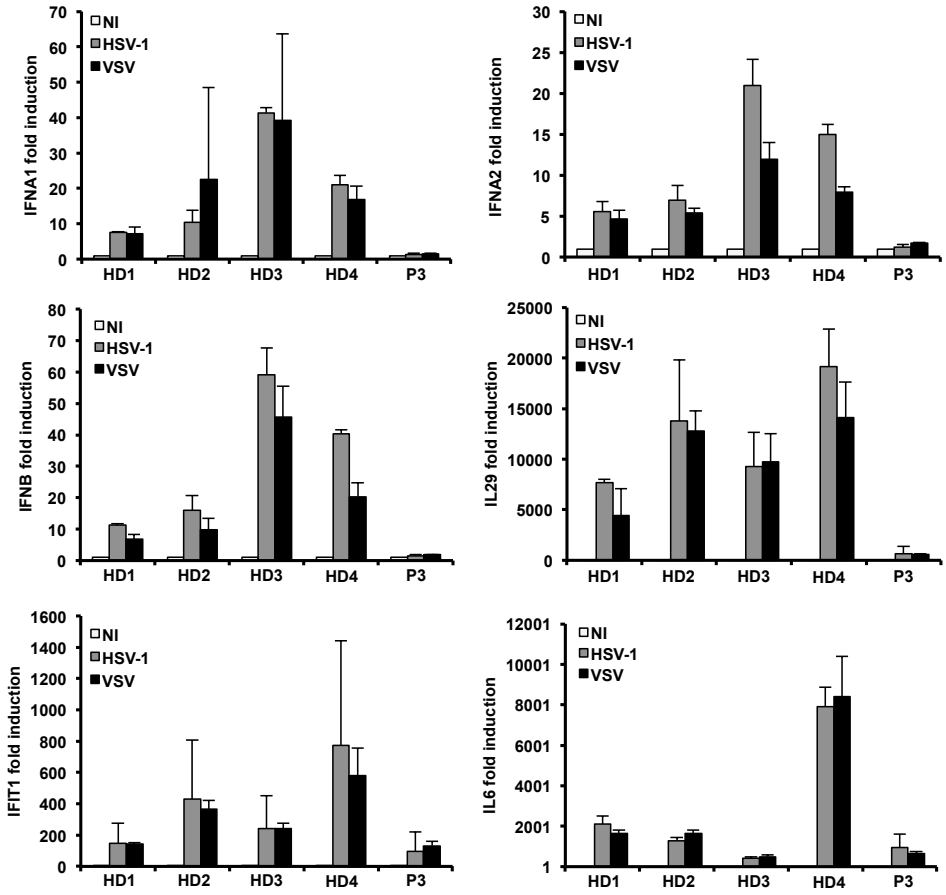


Figure S10. Response to viruses in DOCK2-deficient and healthy donor's PBMCs.

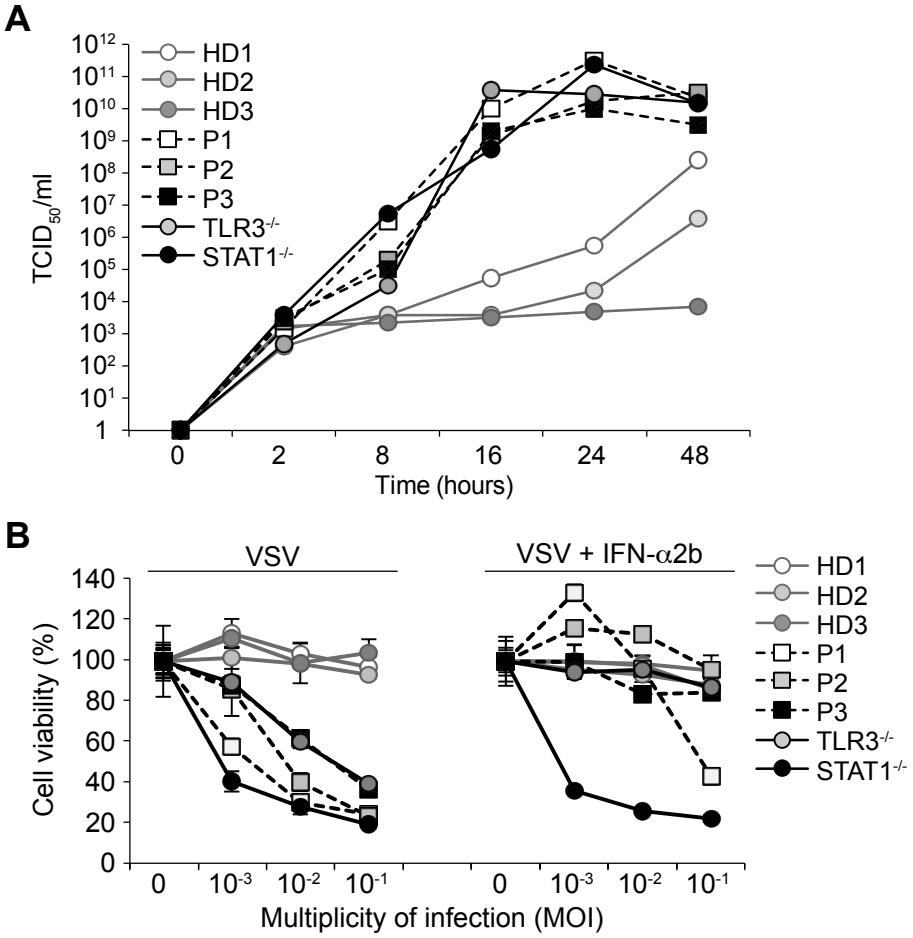
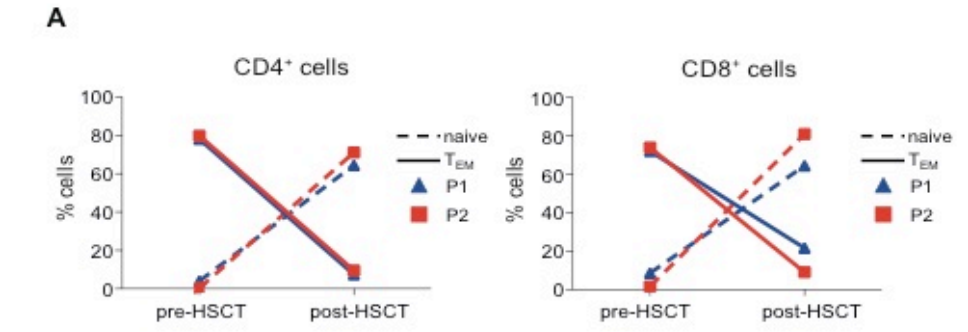


Figure S11. Impaired antiviral interferon responses in of patient-derived DOCK2-deficient fibroblasts. (A) High levels of VSV (vesicular stomatitis virus) replication in DOCK2-deficient SV40-fibroblasts from P1, P2 and P3. $TCID_{50}$ – median tissue culture infective dose. (B) VSV-induced cell death of DOCK2-deficient SV40-fibroblasts (left panel) and the rescue of the VSV-induced cell death by addition of IFN-α2b (right panel).



B

Immune reconstitution post-HSCT in P5

Parameter	pre-HSCT	post-HSCT
ALC (cells/mm³)	2210	4823
CD3⁺ T cells/mm³	830	2878
CD4⁺ T cells/mm³	511	1261
CD8⁺ T cells/mm³	319	1326
CD19⁺ B cells/mm³	1857	610
CD16/56⁺ cells/mm³	1539	1123

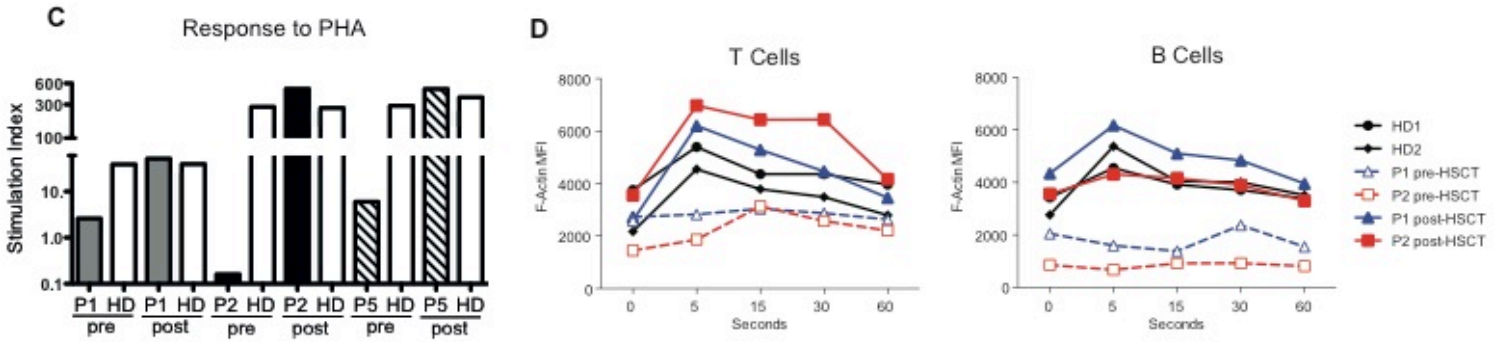


Figure S12. Normalization of T-cell subset distribution, PHA response and actin polymerization after HSCT. (A) Relative numbers of naïve (dashed line) and effector memory (solid line) $CD4^+$ and $CD8^+$ T cells after allogeneic hematopoietic stem cell transplantation in P1 (blue triangles) and P2 (red squares). (B) Absolute cell counts before and HSCT in P5. (C) Recovered response of P1 (grey bars), P2 (filled bars) and P5 (dashed bars) PBMCs to PHA stimulation upon HSCT. (D) Normalization of CXCL12-induced actin polymerization after HSCT in T cells (left) and B cells (right).

Table S1. Homozygous variants identified in patients P1, P3, and P4 and not found in public (dbSNP, 1000 genomes) or in-house databases (upper table). Homozygous variants and heterozygous variants identified in the same gene identified in patient P2 and not found in public (dbSNP, 1000 genomes) or in-house databases (lower table). Non-synonymous homozygous variants and heterozygous variants in the same gene identified in patient P5 and not reported in homozygous form, or present at an allelic frequency greater than 1% among public (ExAC consortium) or in-house databases (lower table).

Patient	Chr.	Start	Reference	Genomic variant in patient	Gene
P1					
	1	38077349	-	A	*RSPO1*
	1	38077420	-	CC	*RSPO1*
	1	54605319	-	C	*CDCP2*
	1	154530774	C	A	*UBE2Q1*
	1	156035733	G	C	*RAB25*
	1	156571221	C	G	*GPATCH4*
	1	160062156	C	T	*IGSF8*
	1	160062977	G	A	*IGSF8*
	1	240255583	CGG	-	*FMN2*
	2	23926150	G	A	*KLHL29*
	2	27278084	A	G	*AGBL5*
	2	29274911	-	AGAGGGGTGGCC	*FAM179A*
	2	29293863	GGG	-	*C2orf71*
	2	42990200	G	A	*OXER1*
	2	88751752	-	G	*FOXI3*
	2	175292599	TTTATCAGTCAAA	-	*SCRN3*
	2	231033633	ATTTTTTTTTTTTTT	ATTAATTTTTTTTTTTTTTT	*SP110*
	2	233351330	C	T	*ECEL1*
	3	16926567	T	C	*PLCL2*
	3	50306757	-	C	*SEMA3B*
	3	53324829	CACTGG	-	*DCP1A*
	3	63898392	-	AGCAGC	*ATXN7*
	3	133969438	-	G	*RYK*
	3	133969490	-	C	*RYK*
	3	141162449	C	T	*ZBTB38*
	4	3076657	GCAGCA	-	*HTT*
	4	155244425	GTTT	-	*DCHS2*
	5	72743299	-	GC	*FOXD1*
	5	167881041	AGG	-	*WWC1*
	5	169122829	A	G	*DOCK2*

5	169468118	-	TA	*DOCK2*
5	176930176	AGG	-	*DOK3*
6	27223069	AGAGCCAGATTAAGG	-	*PRSS16*
6	161519381	CTG	-	*MAP3K4*
7	128587374	GGCCGCCTACTCTGCAGCCGCCCACTCTGC	-	*IRF5*
7	150783917	T	G	*AGAP3*
8	33354255	-	TGA	*MAK16*
8	65494050	-	GCA	*BHLHE22*
8	144511981	TGG	-	*MAFA*
9	79318378	GACAGCCTGCAACGT	-	*PRUNE2*
9	95237066	ATC	-	*ASPN*
9	96439006	-	TGCCTCCACCACACC	*PHF2*
9	123476563	GCGGCG	-	*MEGF9*
10	11374612	CCAGTCTTCCTCTTCGGCATGCCCTGGAAGCTTCTT	-	*CELF2*
10	73499446	A	G	*CDH23*
10	74673110	A	G	*OIT3*
10	74953294	C	G	*FAM149B1*
11	65636053	T	C	*EFEMP2*
12	6777111	CTG	-	*ZNF384*
12	53184620	-	A	*KRT3*
13	72440684	CCGCCG	-	*DACH1*
14	77493761	TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGTTGCTGCTGCTGCTGCTG	TTGCTGCTGCTGCTGCTGCTGCTGTTGCTGCTGCTGCTGCTG	*IRF2BPL*
15	31776229	GGC	-	*OTUD7A*
15	102029688	-	CG	*PCSK6*
16	31099000	C	T	*PRSS53*
16	31508241	G	A	*C16orf58*
16	48120692	A	G	*ABCC12*
16	67241587	C	T	*LRRC29*
17	8384692	C	T	*MYH10*
17	21156532	-	C	*C17orf103*
17	46115074	-	G	*COPZ2*
17	46115123	-	G	*COPZ2*
17	48227384	GC	GGCC	*PPP1R9B*
19	4511548	A	T	*PLIN4*
19	4817312	-	AGG	*TICAM1*
19	8389898	CTC	-	*KANK3*
19	9073059	G	C	*MUC16*
19	14200896	-	G	*SAMD1*
19	18717401	CAG	-	*CRLF1*
19	37879852	CCA	CTGTGCA	*ZNF527*
19	49956468	ACCCCATTCCCATTAGCCCC	-	*ALDH16A1*
19	50832152	T	C	*KCNC3*

	19	52400146	C	T	*ZNF649*
	19	55944675	C	T	*SHISA7*
	20	32664865	-	AGC	*RALY*
	20	35807793	-	ATAGACAGGGCCCCGC GGCCGGCACTCTT	*C20orf132*
	22	20780030	-	C	*SCARF2*
	22	50655681	G	A	*SELO*
P3					
	3	108269959	A	G	*KIAA1524*
	5	112312899	G	C	*DCP2*
	5	140433373	G	A	*PCDHB1*
	5	149215910	A	T	*PPARGC1B*
	5	149772948	G	A	*TCOF1*
	5	150029416	G	A	*SYNPO*
	5	156635924	A	G	*ITK*
	5	169145781	G	T	*DOCK2*
	5	170640698	T	G	*RANBP17*
	9	130635140	C	T	*AK1*
	17	7756658	G	A	*KDM6B*
P4					
	1	248790341	G	A	*OR2T11*
	5	169145753	A	ATG	*DOCK2*
	5	169291354	TCTG	T	*FAM196B*
	9	130117713	CCT	C	*GARNL3*
	9	130164952	C	T	*SLC2A8*
	9	136208374	C	T	*MED22*
	9	139934469	TCC	T	*NPDC1*

Patient	Chr.	Start	Reference	Zygosity	Genomic variant in patient	Gene
P2						
	1	152195728	T	homo	TT	*HRNR*
	2	95847040	-	homo	GCG	*ZNF2*
	2	187559029	-	homo	GCAGCA	*FAM171B*
	3	195452870	G	het	A	*MUC20*
	3	195452872	AG	het	CA	*MUC20*
	4	3076603	-	homo	GCA	*HTT*
	4	147560450	-	homo	TGG	*POU4F2*
	5	169446041	C	het	T	*DOCK2*
	5	169474517	C	het	T	*DOCK2*
	5	171201610	A	het	C	*CTB-78H18.1*
	5	171201615	A	het	C	*CTB-78H18.1*
	9	132630664	GAC	homo	-	*USP20*
	10	50534969	-	homo	ACACACAC	*C10ORF71*
	12	21791410	A	homo	-	*LDHB*
	16	1823389	G	homo	-	*EME2*
P5						
	5	169101338	-	het	GATGTAC	*DOCK2*
	5	169484630	C	het	T	*DOCK2*
	11	74203186	G	homo	A	*LIPT2*
	11	125887165	G	het	A	*CDON*
	11	125891299	C	het	T	*CDON*
	18	29027853	T	het	C	*DSG3*
	18	29054099	C	het	T	*DSG3*

Note that for P5 we have listed all putative compound heterozygous variants, however, it is likely that both identified variants in *CDON, CSHL1* and *DSG3* are paternal alleles since exome sequencing data for the mother of P5 showed none of the two reported variants in each of these genes.

Table S2. Predicted impact of the identified *DOCK2* missense mutations. The mutations were scored using the SIFT, PolyPhen-2 and CADD algorithms. None of the *DOCK2* mutations identified in the affected patients were found in public (dbSNP, 1000 genomes, NHLBI exomes) or our private databases comprising over 2,000 exomes; however, the p.Q1324* allele was present in a large dataset containing >3,500 Finnish exomes at minor allele frequency (MAF) of 0.0002[14] and as heterozygous variant in the ExAc database with a MAF of 0.0001512, the p.R1104W was present as heterozygous variant in the ExAc database with a MAF of 0.00001499 and the p.P1476L was present as heterozygous variant in the ExAc database with a MAF of 0.00009615.

***DOCK2* mutation**	**SIFT score**	**PolyPhen-2 score**	**CADD score**
p.R1104W	0.000	0.983	19.29
p.R751S	0.03	0.991	17.76
p.P1476L	0.000	1.000	32

Supplementary References

1. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 2002;30:97-101.
2. Seelow D, Schuelke M, Hildebrandt F, Nurnberg P. HomozygosityMapper--an interactive approach to homozygosity mapping. Nucleic Acids Res 2009;37:W593-9.
3. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-75.
4. Koskela HL, Eldfors S, Ellonen P, et al. Somatic STAT3 mutations in large granular lymphocytic leukemia. N Engl J Med 2012;366:1905-13.
5. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.
6. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297-303.
7. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009;25:2078-9.
8. Van der Auwera GA, Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K.V., Altshuler, D., Gabriel, S., DePristo, M.A. . From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. Current Protocols in Bioinformatics 2013.
9. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 2003;31:3812-4.
10. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248-9.
11. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46:310-5.
12. Gerstel-Thompson JL, Wilkey JF, Baptiste JC, et al. High-throughput multiplexed T-cell-receptor excision circle quantitative PCR assay with internal controls for detection of severe combined immunodeficiency in population-based newborn screening. Clinical chemistry 2010;56:1466-74.
13. Della Chiesa M, Falco M, Bertaina A, et al. Human cytomegalovirus infection promotes rapid maturation of NK cells expressing activating killer Ig-like receptor in patients transplanted with NKG2C-/- umbilical cord blood. J Immunol 2014;192:1471-9.
14. Lim ET, Wurtz P, Havulinna AS, et al. Distribution and medical impact of loss-of-function variants in the Finnish founder population. PLoS Genet 2014;10:e1004494.